An arithmetic analogue of Fox's triangle removal argument

Pooya Hatami* Sushant Sachdeva† Madhur Tulsiani‡

February 1, 2016

Abstract

We give an arithmetic version of the recent proof of the triangle removal lemma by Fox [Fox11], for the group $\mathbb{F}_2^n$.

A triangle in $\mathbb{F}_2^n$ is a triple (x, y, z) such that $x + y + z = 0$. The triangle removal lemma for $\mathbb{F}_2^n$ states that for every $\varepsilon > 0$ there is a $\delta > 0$, such that if a subset A of $\mathbb{F}_2^n$ requires the removal of at least $\varepsilon \cdot 2^n$ elements to make it triangle-free, then it must contain at least $\delta \cdot 2^{2n}$ triangles. This problem was first studied by Green [Gre05] who proved a lower bound on δ using an arithmetic regularity lemma. Regularity based lower bounds for triangle removal in graphs were recently improved by Fox and we give a direct proof of an analogous improvement for triangle removal in $\mathbb{F}_2^n$.

The improved lower bound was already known to follow (for triangle-removal in all groups), using Fox's removal lemma for directed cycles and a reduction by Král, Serra and Vena [KSV09] (see [Fox11, CF13]). The purpose of this note is to provide a direct Fourier-analytic proof for the group $\mathbb{F}_2^n$.

1 Introduction

The triangle removal lemma for graphs states that for every $\varepsilon > 0$, there exists a $\delta > 0$ such that every graph on n vertices with at most δn^3 triangles can be made

*Institute for Advanced Study, Princeton, NJ, USA. This work was done when the author was a student at University of Chicago. email: `pooyahat@math.ias.edu`. Research supported by the NSF grant No. CCF-1412958.

†Department of Computer Science, Yale University, USA. Part of this work was done when the author was a student at Princeton University, and was visiting TTI Chicago. email: `sachdeva@cs.yale.edu`

‡Toyota Technological Institute at Chicago. email: `madhurt@ttic.edu`. Research supported by NSF Career Award CCF-1254044.

triangle-free by deleting less than εn^2 edges. Contrapositively, this means that if a graph is at least ε-far from being triangle-free, *i.e.*, one needs to delete more than εn^2 edges to make it triangle free, then it must have at least δn^3 triangles.

The lemma was originally proved by Ruzsa and Szemerédi [RS78] with a bound of $\delta \geq 1/\mathsf{Tower}(\mathrm{poly}(1/\varepsilon))$, where $\mathsf{Tower}(i)$ denotes a tower of twos of height i, *i.e.*, $\mathsf{Tower}(i) = 2^{\mathsf{Tower}(i-1)}$ for $i \geq 1$, and $\mathsf{Tower}(0) = 1$. It took over three decades for this bound to be improved to $\delta \geq 1/\mathsf{Tower}(O(\log(1/\varepsilon)))$ in a remarkable paper of Fox [Fox11].

The above lemma has a direct application to property testing as it states that for a graph which is ε-far from being triangle-free, a random triple of vertices is guaranteed to form a triangle with probability at least δ. Thus if we test $\Omega(1/\delta)$ random triples of vertices, we will find a triangle with constant probability. This test can then distinguish such a graph from one which is triangle-free, since, for a triangle-free graph, the probability of the test succeeding is 0. The triangle removal lemma and its generalizations to removal of more general graphs (instead of a triangle) have several interesting applications in mathematics, and we refer the reader to the survey [CF13] for a detailed discussion.

An arithmetic version of the triangle removal problem was considered by Green [Gre05]. Let G be an Abelian group with $|G| = N$ and let $A \subseteq G$ be an arbitrary subset. We call a triple $(x, y, z) \in A^3$ a *triangle* if $x + y + z = 0$. Similar to the graph case, we say that A is ε-far from being triangle-free if one needs to remove at least εN elements from A to make it triangle-free. Green proved an arithmetic analogue of Szemerédi's regularity lemma [Sze75], used in the proof of Ruzsa and Szemerédi, and proved a triangle removal lemma for Abelian groups.

Theorem 1.1 (Green [Gre05]) *For all $\varepsilon \in (0, 1]$, there exists a $\delta \geq 1/\mathsf{Tower}(\mathrm{poly}(1/\varepsilon))$ such that for an Abelian group G with $|G| = N$, if a subset $A \subseteq G$ is ε-far from being triangle-free, then A must contain at least δN^2 triangles.*

Other than being useful for proving the above result, Green's arithmetic analogue of Szemerédi's regularity lemma has had many other applications in combinatorics [GN14a, GN14b] and computer science [KO09, BGS15, BGRS12]. In addition to leading to interesting analogues of combinatorial statements, the arithmetic setting has at least one more advantage: the proofs of these statements often proceed by partitioning the underlying spaces, and because of the structure provided by subgroups and cosets, the arguments are often cleaner (specially for vector spaces over finite fields). This makes the arithmetic setting more attractive in the search for quantitative improvements to the results.

In this paper, we present an arithmetic analogue of the proof by Fox, for the group $\mathbb{F}_2^n$. The argument can also be extended to other Abelian groups, but we restrict ourselves to $\mathbb{F}_2^n$ for simplicity.

Theorem 1.2 *For all $\varepsilon \in (0,1]$, there exists a $\delta \geq 1/\mathsf{Tower}(O(\log((1/\varepsilon)))$ such that for all $n \in \mathbb{N}$ and $N \stackrel{\text{def}}{=} 2^n$, any subset $A \subseteq \mathbb{F}_2^n$ which is ε-far from being triangle-free, must contain at least δN^2 triangles.*

We remark that the above result (for all groups) already follows from a version of the removal lemma for directed cycles, using a reduction by Král, Serra and Vena [KSV09]. This was already observed by Fox [Fox11] (see also [CF13]). However, we believe some of the Fourier analytic notions that come up in a direct arithmetic proof might be of independent interest. Also, a direct proof makes it somewhat more transparent how the argument partitions the underlying group, which might be useful for further improvements. We present a sketch of the proof below.

Known lower bounds for triangle removal. The original proof by Ruzsa and Szemerédi used Szemerédi's regularity lemma [Sze75]. However, the known lower bounds for the regularity lemma and its variants (see [Gow97] and [CF13]) imply that this approach necessarily obtains a bound on $1/\delta$ which is at least $\mathsf{Tower}(\mathrm{poly}(1/\varepsilon))$. Fox's argument [Fox11] manages to avoid using the regularity lemma directly (although his proof still follows the same outline as the proof of the regularity lemma), thus obtaining $1/\delta \leq \mathsf{Tower}(O(\log(1/\varepsilon)))$.

In terms of lower bounds on $1/\delta$, it was shown by Alon [Alo02] that one must have $1/\delta \geq 2^{\Omega(\log^2(1/\varepsilon))}$ for triangle removal in graphs. For the problem of triangle removal in groups, in the case when the group is $\mathbb{Z}/p\mathbb{Z}$, Green [Gre05] gives a similar lower bound of $1/\delta \geq 2^{\Omega(\log^2(1/\varepsilon))}$. However, in the case of $\mathbb{F}_2^n$, the only known lower bounds are polynomial in ε. Bhattacharyya and Xie [BX10] show that for triangle removal in $\mathbb{F}_2^n$ one must have $1/\delta \geq (1/\varepsilon)^{8.487}$, which was later improved to $(1/\varepsilon)^{13.239}$ by [FK14]. These lower bounds remain quite far from the known upper bounds on $1/\delta$.

1.1 Proof Sketch

We will give a proof by contradiction. Let G denote the group $\mathbb{F}_2^n$, and assume, to the contrary, that we have a set $A \subseteq G$ that is ε-far from being triangle-free, and has at most $\delta|G|^2$ triangles. We work with a subset $A' \subseteq A$ that is the union of a maximal collection of element disjoint triangles, *i.e.* no two triples representing triangles share a common element. Since $A \backslash A'$ is triangle-free, and A is ε-far from being triangle-free, $|A'| \geq \varepsilon|G|$. Define $\varepsilon_0 \stackrel{\text{def}}{=} {}^{|A'|}/_{|G|}$.

In the rest of the sketch, we denote A' by A, and ε_0 by ε. The proof is based on a potential-increment argument. At every step in the proof, we have a partition of G into T cosets of a subgroup $H \preceq G$ where $T \stackrel{\text{def}}{=} |{}^{G}/_{H}|$, and hence an induced partition of A according to cosets of H. We measure the *mean entropy* of the partition, defined as $\mathbb{E}_g\left[\mathsf{Ent}\left(\frac{|A\cap(H+g)|}{|H|}\right)\right]$, where $\mathsf{Ent}(x) \stackrel{\text{def}}{=} x\log x$ for $x \in (0,1]$ and $\mathsf{Ent}(0) \stackrel{\text{def}}{=} 0$. Observe that the mean entropy is always non-positive, and by convexity it is at least $\varepsilon\log\varepsilon$. Our main lemma proves that if δ is much smaller than $\varepsilon^3 \cdot {}^{|G|^2}/_{T^2}$, then we can *shatter* A, *i.e.*, the current partition of A can be refined according to cosets of $H' \preceq H$, such that the mean entropy increases by $\Omega(\varepsilon)$, and $|{}^{G}/_{H'}| \leq 2^{T\cdot O(\varepsilon^{-3})}$. In essence, the size of the partition, and hence the bound on δ, are one exponential larger at every step. Since the mean entropy is always non-positive, this process must stop after $O(\log\varepsilon^{-1})$ rounds, giving the required bound on δ.

In order to show that we can shatter A if it has too few triangles, it is convenient to equate A with its indicator function $A : G \to \{0,1\}$, and the number of triangles with the sum $\sum_{x+y+z=0} A(x)A(y)A(z)$. Suppose we want to count the number of triangles between two cosets of H, viz. $H+g_1$ and $H+g_2$, and a third coset of H', $H'+g_3+z$, where $g_1+g_2+g_3=0$. Assume A has density ε on all the three cosets. As a thought experiment, if A was the constant function ε on $H'+g_3+z$, then the "number of triangles" between the three cosets is given by $\varepsilon|H||H'|\,\mathbb{E}_{g\in H}\left[\frac{|A\cap(H'+g+g_1)|}{|H'|}\cdot\frac{|A\cap(H'+g+g_2+z)|}{|H'|}\right]$. If this is significantly smaller than $\varepsilon^3|H||H'|$, then a Markov argument implies that partitioning according to cosets of H' *shatters* A *i.e.*, a constant fraction of the cosets have density significantly smaller than ε. A defect version of Jensen's inequality then implies that the mean entropy of the new partition is larger by $\Omega(\varepsilon)$.

Of course, A is not necessarily a constant function. However, a very similar argument works if all the non-zero Fourier coefficients of the function A on $H'+g_3+z$ are much smaller than ε in absolute value – we call such functions *superregular*, in analogy with the proof from [Fox11]. The last part of the proof is to find a *superregular decomposition* of A on $H+g_3$, *i.e.* approximating $A\cap(H+g_3)$ by a sum of superregular functions. If A is not superregular on $H+g_3$, we pick a large Fourier coefficient $\eta \in \widehat{H}$, partition $x \in H$ according to the value of $\langle x,\eta\rangle$, and restrict ourselves to the part with greater density. If this part is not superregular, we repeat this procedure. Since the density on any part is at most 1, this process must end with a superregular part. We remove this set from A and repeat the procedure until most of A is covered, to find the required superregular decomposition.

This completes the proof sketch of Theorem 1.2.

Analogies and differences with the proof in [Fox11]. The proof of our main theorem on triangle-removal in groups is analogous to the proof of triangle-removal in graphs from the work of Fox [Fox11]. Nevertheless, we need to give the appropriate arithmetic analogues of the definitions and proofs. Though the proof in this paper is self-contained, for the readers who are familiar with the work of Fox [Fox11], we point out the analogies and the differences between the two proofs.

At every step in our proof, as in [Fox11], we have a partition of the underlying set, the group G in our case. However, our partition is structured, and consists of all the cosets of one fixed subgroup, compared to an arbitrary partition in [Fox11]. Our definitions of the potential function and shattering are similar, except that we use the densities of the cosets, instead of the edge densities between pairs of subsets used in [Fox11]. Our notion of regularity is quite different from that in [Fox11], and is based on Fourier coefficients, similar to the one used in regularity lemmas for Abelian groups [Gre05]. The superregular decomposition that we find for a given set is analogous to the collection of superregular tuples in a graph, constructed in [Fox11].

2 Preliminaries

Fix a positive integer n. Throughout the paper, we denote $G \stackrel{\text{def}}{=} \mathbb{F}_2^n$ and $N \stackrel{\text{def}}{=} |G| = 2^n$. The notation $H \preceq G$ denotes that H is a subgroup of G. We use $\widehat{H}$ to denote the dual group of H. Denote by $H^{\perp}$ coset group of H in G, which we also use to denote a set of coset representatives. In some cases which will be clear from the context, by abuse of notation, we use the common definition of $H^{\perp} = \{y \mid \sum_{i=1}^{n} x_i \cdot y_i = 0 \ \ \forall x \in H\}$. Note that for $G = \mathbb{F}_2^n$,

$$\hat{H} \;\cong\; G/H^{\perp} \;\cong\; H\,,$$

and hence $H^{\perp}$ can also be thought of as the coset group G/H.

Given a set $A \subseteq G$, three elements $x, y, z \in A$ are said to form a *triangle* if $x + y + z = 0$. A is said to be ε-far from being triangle-free, if at least εN elements need to be removed from A in order that it contains no triangles.

We abuse notation and use A to denote both the set A, and also its characteristic function $A : G \to \{0, 1\}$. Given a subgroup $H \preceq G$, and an element $g \in G$, we define $A_H^g : H \to \{0, 1\}$ as

$$A_H^g(x) \stackrel{\text{def}}{=} A(x + g).$$

Let $\mathbb{E}_{x \in H}[\cdot]$ denote the expectation when x is drawn uniformly from H. For a function $f : H \to \mathbb{R}$, we define its Fourier coefficients as follows: For $\eta \in \widehat{H}$, define

$$\widehat{f}(\eta) \stackrel{\text{def}}{=} \mathop{\mathbb{E}}_{x \in H} \left[f(x) \chi_{\eta}(x) \right].$$

Jensen's inequality states that if Ent is a convex function, $\varepsilon_1, \ldots, \varepsilon_s$ are nonnegative real numbers such that $\sum_{i=1}^{s} \varepsilon_i = 1$, then, for any real numbers $x_1, \ldots, x_s$,

$$\varepsilon_1 \mathsf{Ent}(x_1) + \cdots + \varepsilon_s \mathsf{Ent}(x_s) \geq \mathsf{Ent}(\varepsilon_1 x_1 + \cdots + \varepsilon_s x_s). \qquad (1)$$

Jensen's inequality immediately implies the following simple lemma, which will be useful later.

Lemma 2.1 (Lemma 6, [Fox11]) *Let* $\mathsf{Ent} : \mathbb{R}_{\geq 0} \to \mathbb{R}$ *be a convex function,* $\varepsilon_1, \ldots, \varepsilon_s$ *and* $x_1, \ldots, x_s$ *be nonnegative real numbers with* $\sum_{i \in [s]} \varepsilon_i = 1$*. For* $I \subseteq [s]$*, let* $c = \sum_{i \in I} \varepsilon_i$*,* $u = \sum_{i \in I} \varepsilon_i x_i / c$*, and* $v = \sum_{i \in [s] \setminus I} \varepsilon_i x_i / (1 - c)$*. Then we have*

$$\sum_{i \in [s]} \varepsilon_i \mathsf{Ent}(x_i) \geq c \mathsf{Ent}(u) + (1 - c) \mathsf{Ent}(v).$$

2.1 Entropy

Our proof of the main theorem will be based on a potential-increment argument, where our potential function will be the *mean entropy of a partition*, as defined below.

Definition 2.2 (Entropy) *Define the* entropy function $\mathsf{Ent} : \mathbb{R}_{\geq 0} \to \mathbb{R}$ *as* $\mathsf{Ent}(x) \stackrel{\text{def}}{=} x \log x$*, for* $x \in \mathbb{R}_{>0}$*, and* $\mathsf{Ent}(0) = 0$*.*

Definition 2.3 (Mean Entropy of a Partition) *Let* $H' \preceq H \preceq G$*, and* $g \in G$*. Given a set* $A \subseteq G$*, if we partition the coset* $H + g$ *as cosets of* H'*, define the mean entropy of this partition as follows*

$$\mathsf{Ent}_A(H + g, H') \stackrel{\text{def}}{=} \mathop{\mathbb{E}}_{g_1 \in H + g} \mathsf{Ent}\left(\frac{|A \cap (H' + g_1)|}{|H'|} \right) = \mathop{\mathbb{E}}_{g_1 \in H} \mathsf{Ent}\left(\frac{|A \cap (H' + g + g_1)|}{|H'|} \right).$$

Analogously, the mean entropy of partitioning A *on* G *according to cosets of* $H \preceq G$ *is defined as*

$$\mathsf{Ent}_A(G, H) \stackrel{\text{def}}{=} \mathop{\mathbb{E}}_{g_1 \in G} \mathsf{Ent}\left(\frac{|A \cap (H + g_1)|}{|H|} \right).$$

Remark 2.4 *Assume* $A \subseteq G$ *has been partitioned according to cosets of a subgroup* $H \preceq G$*. For a refinement to this partition according to* $H' \preceq H$ *we have*

$$\mathsf{Ent}_A(G, H') = \mathop{\mathbb{E}}_{g \in G} \mathsf{Ent}\left(\frac{|A \cap (H' + g)|}{|H'|} \right) = \mathop{\mathbb{E}}_{g \in G} \mathsf{Ent}_A(H + g, H').$$

In order to quantify the increase in the mean entropy because of a shattering, we need the following defect version of Jensen's inequality for the entropy function. We include a proof for completeness.

Lemma 2.5 (Defect inequality for entropy. [Fox11], Lemma 7) *Let* $\varepsilon_1, \ldots, \varepsilon_s$*, and* $x_1, \ldots, x_s$*, be nonnegative real numbers with* $\sum_{i \in [s]} \varepsilon_i = 1$*, and* $a = \sum_{i \in [s]} \varepsilon_i x_i$*. Suppose* $\beta < 1$*, and* $I \subseteq [s]$ *is such that* $x_i \leq \beta a$ *for all* $i \in I$*. Let* $c = \sum_{i \in I} \varepsilon_i$*. Then,*

$$\sum_{i \in [s]} \varepsilon_i \mathsf{Ent}(x_i) \geq \mathsf{Ent}(a) + (1 - \beta + \mathsf{Ent}(\beta)) ca.$$

Proof: We know that $c < 1$ since otherwise, $a = \sum_{1\leq i\leq s} \varepsilon_i x_i = \sum_{i\in I} \varepsilon_i x_i \leq \beta a < a$, a contradiction. The cases when a or c are equal to 0 follow immediately from Jensen's inequality (Equation (1)), therefore we may assume that $a, c \neq 0$.

Letting $u \stackrel{\text{def}}{=} \frac{1}{c}\sum_{i\in I} \varepsilon_i x_i$ and $v \stackrel{\text{def}}{=} \frac{1}{1-c}\sum_{i\notin I} \varepsilon_i x_i$, we have,

$$\begin{aligned}
\sum_{1\leq i\leq s} \varepsilon_i \mathsf{Ent}(x_i) &\geq c\mathsf{Ent}(u) + (1-c)\mathsf{Ent}(v) \\
&= \mathsf{Ent}(a) + ca\mathsf{Ent}(u/a) + (1-c)a\mathsf{Ent}(v/a) \\
&\geq \mathsf{Ent}(a) + ca\mathsf{Ent}(u/a) + ca(1-u/a) \\
&= \mathsf{Ent}(a) + \big(\mathsf{Ent}(u/a) + 1 - u/a\big)ca \\
&\geq \mathsf{Ent}(a) + (1-\beta+\mathsf{Ent}(\beta))ca,
\end{aligned}$$

where the first inequality is from Lemma 2.1, the first equality follows from the definition of the entropy function Ent, the second inequality follows from the fact that $\mathsf{Ent}(\frac{v}{a}) = \mathsf{Ent}(\frac{1-uc/a}{1-c}) = \mathsf{Ent}(1+\frac{(1-u/a)c}{1-c}) > \frac{(1-u/a)c}{1-c}$, and the last inequality follows from the fact that $u/a \leq \beta$ and that $\mathsf{Ent}(x) + 1 - x$ is a decreasing function on the interval $[0,1]$. ∎

The following lemma shows how the mean entropy of a partition compares to that of its refinement.

Lemma 2.6 (Entropy - Basic inequalities) *Let $H' \preceq H \preceq G$, and $A \subseteq G$.*

1. $0 \geq \mathsf{Ent}_A(G,H) \geq \mathsf{Ent}_A(G,G) = \mathsf{Ent}\left(\frac{|A|}{|G|}\right)$.
2. *For every $g \in G$, we have $\mathsf{Ent}_A(H+g,H') \geq \mathsf{Ent}_A(H+g,H)$, and therefore $\mathsf{Ent}_A(G,H') \geq \mathsf{Ent}_A(G,H)$.*

Proof: Both parts follow from convexity of Ent, and Jensen's inequality (Equation (1)). ∎

In Lemma 4.2, we will show how shattering a partition can substantially increase its mean entropy, which will allow us to use the mean entropy as a potential function.

3 Super-regularity

In this section, we define a notion of regularity, and show how one can approximate any set A as a sum of regular parts.

Definition 3.1 (ρ-Superregularity) *Let $H \preceq G$, and $g \in G$. Given a function $f : H+g \to \mathbb{R}$, say that f is ρ-superregular on $H+g$ if, for every $\eta \in \widehat{H}$, $\eta \neq 0$, we have, $\left|\widehat{f_H^g}(\eta)\right| \leq \rho \cdot \widehat{f_H^g}(0)$, where the function $f_H^g : H \to \mathbb{R}$ is defined as $f_H^g(x) \stackrel{\text{def}}{=} f(x+g)$ for all $x \in H$.*

Similar notions of regularity have been used in the proof of a Szemerédi type regularity lemma for $\mathbb{F}_2^n$ [Gre05], and are well-studied as notions of pseudorandomness for subsets of abelian groups (*e.g.* see [RCLG92, Gow98]).

Given a set A, the following lemma identifies a subset of A that is superregular.

Lemma 3.2 (Finding Superregular Parts) *Let H be a subgroup of G. Given an element $g \in G$, a desired regularity parameter $\rho \in (0,1]$, a density parameter $d > 0$, and a set $A \subseteq H+g$ such that $|A| \geq d|H|$, we can find a triple (A_1, H_1, z_1) such that:*

1. *H_1 is a subgroup of H satisfying $|H/H_1| \leq 2^{\log_{(1+\rho)}(1/d)}$.*
2. *$z_1 \in H \cap H_1^{\perp}$.*
3. *$A_1 \subseteq A$, $A_1 \subseteq H_1 + z_1 + g$, and $|A_1| \geq d|H_1|$.*
4. *The indicator function of A_1 restricted to $H_1 + z_1 + g$ is ρ-superregular on $H_1 + z_1 + g$.*

Proof: We give an iterative procedure to find A_1, H_1 and z_1. Initialize $A_1 \stackrel{\text{def}}{=} A$, $H_1 \stackrel{\text{def}}{=} H$ and $z_1 = 0$. Observe that $z_1 \in H \cap H_1^{\perp}$, and $|A_1| \geq d\,|H_1|$.

If the indicator function of A_1 restricted to $H_1 + g + z_1$ is ρ-superregular on $H_1 + g + z_1$, we are done and we can return (A_1, H_1, z_1).

Otherwise, we must have $\eta \in \widehat{H_1} \backslash \{0\}$ for which $|\widehat{{A_1}_{H_1}^{g+z_1}}(\eta)| \geq \rho \cdot \widehat{{A_1}_{H_1}^{g+z_1}}(0)$. Define $H_2 \stackrel{\text{def}}{=} H_1 \cap \{h \mid \langle h, \eta \rangle = 0\}$, so that $|H_1/H_2| = 2$.

If $|A_1 \cap (H_2 + g + z_1)| \geq |A_1 \cap (H_2 + g + z_1 + \eta)|$, let $z_2 \stackrel{\text{def}}{=} z_1$. Otherwise, let $z_2 \stackrel{\text{def}}{=} z_1 + \eta$. Note that $z_2 \in H \cap H_2^{\perp}$.

Defining $A_2 \stackrel{\text{def}}{=} A_1 \cap (H_2 + g + z_2)$, we have

$$\begin{aligned}
|\widehat{{A_1}_{H_1}^{g+z_1}}(\eta)| &= \frac{|A_2| - |A_1 \cap (H_2 + g + z_2 + \eta)|}{|H_1|} \\
&\geq \rho \widehat{{A_1}_{H_1}^{g+z_1}}(0) = \rho \cdot \frac{|A_1|}{|H_1|},
\end{aligned}$$

moreover $\frac{|A_2| + |A_1 \cap (H_2 + g + z_2 + \eta)|}{|H_1|} = \frac{|A_1|}{|H_1|}$. Consequently,

$$2 \cdot \frac{|A_2|}{|H_1|} \geq (1+\rho) \cdot \frac{|A_1|}{|H_1|}.$$

Thus since $|H_1/H_2| = 2$, we have that $\frac{|A_2|}{|H_2|} \geq (1+\rho) \cdot \frac{|A_1|}{|H_1|}$, and $|A_2| \geq d|H_2|$ in particular.

Now, we let $H_1 \stackrel{\text{def}}{=} H_2$, $A_1 \stackrel{\text{def}}{=} A_2$ and $z_1 \stackrel{\text{def}}{=} z_2$, and repeat the whole procedure. At every step, the triple (A_1, H_1, z_1) satisfies properties 2 and 3.

Since the density of A_1, *i.e.* $\frac{|A_1|}{|H_1|}$, is increasing by a factor of $(1+\rho)$ at every step, and can be at most 1, in $\log_{(1+\rho)}(1/d)$ iterations we must find a triple such that the restriction of A_1 to $H_1 + z_1 + g$ is ρ-superregular on $H_1 + z_1 + g$. Thus, the final triple satisfies property 4. Finally, property 1 also holds since at every step, $|H/H_1|$ increases by a factor of 2. ∎

Repeatedly applying the above lemma gives us the following corollary, which allows us to decompose a set into superregular parts.

Corollary 3.3 (Superregular Decomposition) *Let H be a subgroup of G. Given an element $g \in G$, a desired regularity parameter $\rho \in (0,1]$, a density parameter $d \in (0,1]$, and a set $A \subseteq H + g$, we can find a positive integer t, and a collection of t triples (A_i, H_i, z_i), $i = 1, \ldots, t$, such that $\{A_i\}_{i=1}^{t}$ are disjoint subsets of A satisfying $|A \backslash (A_1 \cup \cdots \cup A_t)| \leq d|H|$, and for every $i = 1, \ldots, t$:*

1. *H_i is a subgroup of H satisfying $|H/H_i| \leq 2^{\log_{(1+\rho)}(1/d)}$.*
2. *$z_i \in H \cap H_i^{\perp}$.*
3. *$A_i \subseteq A$, $A_i \subseteq H_i + z_i + g$, and $|A_i| \geq d\,|H_i|$.*
4. *The indicator function of A_i restricted to $H_i + z_i + g$ is ρ-superregular on $H_i + z_i + g$.*

Proof: We build the collection of triples one at a time. At the i^{th} step, let $B \stackrel{\text{def}}{=} A \backslash (A_1 \cup \cdots \cup A_{i-1})$. If $|B| \leq d|H|$, we can return the triples found till step $(i-1)$.

So we can assume $|B| > d|H|$, and apply Lemma 3.2 with subgroup H, element g, regularity parameter ρ, density parameter d, and subset $B \subseteq H + g$, to find (A_i, H_i, z_i) satisfying the required properties 1-4. Moreover, since $A_i \subseteq A \backslash (A_1 \cup \ldots \cup A_{i-1})$, A_i is disjoint from $A_1, \ldots, A_{i-1}$.

We add the triple (A_i, H_i, z_i) to the collection and iterate. ∎

4 Shattering

In this section, we prove that if a set A contains few triangles from three cosets, then we can partition at least one of the cosets into parts with significantly varying densities. In order to state the concerned lemma formally, we need to define *shattering*. The following definition is similar to that used in Fox's proof of the graph removal lemma [Fox11].

Definition 4.1 (Shattering) *Let H be a subgroup of G. Given $g \in G$ and a set $A \subseteq G$, define $d \stackrel{\text{def}}{=} \frac{|A \cap (H+g)|}{|H|}$. Given a subgroup $H' \preceq H$, and parameters $\alpha, \beta \in (0,1]$, we say that H' (α, β, k)-shatters A on $H+g$ if:*

1. $|{}^{H}/_{H'}| \leq 2^k$, *and*
2. $\mathbb{P}_{g' \in H}\left[|A \cap (H' + g + g')| \leq \beta d\,|H'|\right] \geq \alpha$.

Namely, partitioning $H+g$ according to cosets of H' results in significant variation in the density of A.

Once we find a shattering, the defect version of Jensen's inequality allows us to prove that the mean entropy of the partition increases.

Lemma 4.2 (Entropy Increment) *Let $H' \preceq H \preceq G$, and $A \subseteq G$. Let $g \in G$. If H' (α, β, k)-shatters A on $H+g$ then*

$$\mathsf{Ent}_A(H+g, H') \geq \mathsf{Ent}_A(H+g, H) + (1 - \beta + \mathsf{Ent}(\beta))\alpha \cdot \frac{|A \cap (H+g)|}{|H|}.$$

Proof: Follows from the definitions of Ent_A and shattering, and applying Lemma 2.5. ■

Now, we can state the main lemma of this section.

Lemma 4.3 (Shattering Lemma) *Let $H \preceq G$, and $g_1, g_2, g_3 \in G$ such that $g_1 + g_2 + g_3 = 0$. We are given a set $A \subseteq G$ with densities d_1, d_2, and d_3 on $H+g_1$, $H+g_2$, and $H+g_3$ respectively. Then, either A contains at least $\frac{1}{8}d_1 d_2 d_3 |H|^2$ triangles, or else, there is a subgroup $H' \preceq H$ such that H' $\left({}^{1}/_{20}, {}^{3}/_{4}, \log_{1+\rho}\left({}^{2}/_{d_1}\right)\right)$-shatters A on at least one of $H+g_2$ or $H+g_3$, where $\rho = \frac{d_2 d_3}{4}$.*

We first prove some necessary lemmas, and then give a proof of the above lemma at the end of this section. The following lemma allows us to count the number of triangles between three different sets.

Lemma 4.4 (Triangle Counting) *Let $H' \preceq H \preceq G$. We are given $g_1, g_2, g_3 \in G$ such that $g_1 + g_2 + g_3 = 0$, and $z_1 \in H$. For any sets $A, B, C \subseteq G$, the number of triangles between $A \cap (H' + g_1 + z_1), B \cap (H + g_2)$, and $C \cap (H + g_3)$ is given by,*

$$|H||H'| \sum_{\alpha \in \widehat{H'}, \eta \in \widehat{H'}^{\perp} \cap \widehat{H}} \widehat{A_{H'}^{g_1+z_1}}(\alpha)\widehat{B_H^{g_2}}(\alpha+\eta)\widehat{C_H^{g_3}}(\alpha+\eta)\chi_{\alpha+\eta}(z_1).$$

Proof: We first observe that for any $x_1 \in H'$, and $x_2, x_3 \in H$, three elements $x_1 + g_1 + z_1 \in A$, $x_2 + g_2 \in B$, and $x_3 + g_3 \in C$, form a triangle iff $x_1 + x_2 + x_3 = z_1$, since,

$$(x_1 + g_1 + z_1) + (x_2 + g_2) + (x_3 + g_3) = x_1 + x_2 + x_3 + z_1.$$

Thus, in order to count the triangles in the three cosets, we count all such triples x_1, x_2, x_3. The number of triangles equals:

$$\begin{aligned}
&\sum_{\substack{x_1 \in H', x_2, x_3 \in H \\ x_1 + x_2 + x_3 = z_1}} A(x_1 + g_1 + z_1) B(x_2 + g_2) C(x_3 + g_3) \\
&= \sum_{\substack{x_1 \in H', x_2, x_3 \in H \\ x_1 + x_2 + x_3 = z_1}} A_{H'}^{g_1 + z_1}(x_1) B_H^{g_2}(x_2) C_H^{g_3}(x_3) \\
&= \sum_{\substack{x_1 \in H', x_2, x_3 \in H \\ x_1 + x_2 + x_3 = z_1}} \sum_{\alpha \in \widehat{H'}, \beta, \gamma \in \widehat{H}} \widehat{A_{H'}^{g_1 + z_1}}(\alpha) \widehat{B_H^{g_2}}(\beta) \widehat{C_H^{g_3}}(\gamma) \chi_\alpha(x_1) \chi_\beta(x_2) \chi_\gamma(x_3) \\
&= \sum_{x_1 \in H', x_2 \in H} \sum_{\alpha \in \widehat{H'}, \beta, \gamma \in \widehat{H}} \widehat{A_{H'}^{g_1 + z_1}}(\alpha) \widehat{B_H^{g_2}}(\beta) \widehat{C_H^{g_3}}(\gamma) \chi_\alpha(x_1) \chi_\beta(x_2) \chi_\gamma(x_1 + x_2 + z_1) \\
&= |H||H'| \sum_{\alpha \in \widehat{H'}, \beta, \gamma \in \widehat{H}} \widehat{A_{H'}^{g_1 + z_1}}(\alpha) \widehat{B_H^{g_2}}(\beta) \widehat{C_H^{g_3}}(\gamma) \mathop{\mathbb{E}}_{\substack{x_1 \in H' \\ x_2 \in H}} \left[\chi_{\alpha+\gamma}(x_1) \chi_{\beta+\gamma}(x_2) \chi_\gamma(z_1) \right] \\
&= |H||H'| \sum_{\alpha \in \widehat{H'}, \beta \in \widehat{H}} \widehat{A_{H'}^{g_1 + z_1}}(\alpha) \widehat{B_H^{g_2}}(\beta) \widehat{C_H^{g_3}}(\beta) \mathop{\mathbb{E}}_{x_1 \in H'} \left[\chi_{\alpha+\beta}(x_1) \chi_\beta(z_1) \right] \\
&= |H||H'| \sum_{\alpha \in \widehat{H'}, \beta \in \widehat{H}, \alpha+\beta \in \widehat{H'}^{\perp}} \widehat{A_{H'}^{g_1 + z_1}}(\alpha) \widehat{B_H^{g_2}}(\beta) \widehat{C_H^{g_3}}(\beta) \chi_\beta(z_1) \\
&= |H||H'| \sum_{\alpha \in \widehat{H'}, \eta \in \widehat{H'}^{\perp} \cap \widehat{H},} \widehat{A_{H'}^{g_1 + z_1}}(\alpha) \widehat{B_H^{g_2}}(\alpha + \eta) \widehat{C_H^{g_3}}(\alpha + \eta) \chi_{\alpha+\eta}(z_1),
\end{aligned}$$

where the last equality follows by viewing $\widehat{H'}$ as a subgroup of $\widehat{H}$. ■

Next, we use the above lemma to show that if the number of triangles between three sets is small, then we can shatter at least one of them.

Lemma 4.5 (Shattering with a Superregular Part) *Let $H' \preceq H \preceq G$. We have $g_1, g_2, g_3 \in G$ such that $g_1 + g_2 + g_3 = 0$, and $z_1 \in H$. Suppose we are given three sets $A, B, C \subseteq G$, such that*

$$\frac{|A \cap (H' + g_1 + z_1)|}{|H'|} = d_1, \quad \frac{|B \cap (H + g_2)|}{|H|} = d_2, \quad \frac{|C \cap (H + g_3)|}{|H|} = d_3.$$

Also assume that the indicator function of A restricted to $H' + g_1 + z_1$ is $\frac{d_2 d_3}{4}$-superregular on $H' + g_1 + z_1$.

Then, either there are at least $\frac{1}{4}d_1d_2d_3|H||H'|$ triangles between $A \cap (H' + g_1 + z_1)$, $B \cap (H + g_2)$, and $C \cap (H + g_3)$, or else, H' $(1/20, 3/4, \log_2 |H/H'|)$-shatters either B on $H + g_2$, or C on $H + g_3$.

Proof: We first use Lemma 4.4 to count the triangles between $A \cap (H' + g_1 + z_1)$, $B \cap (H + g_2)$, and $C \cap (H + g_3)$.

$$
\begin{aligned}
\widehat{H'}^{\text{No. of triangles}} &= |H||H'| \sum_{\alpha \in \widehat{H'}, \eta \in \widehat{H'}^{\perp} \cap \widehat{H}} \widehat{A_{H'}^{g_1+z_1}}(\alpha)\widehat{B_H^{g_2}}(\alpha+\eta)\widehat{C_H^{g_3}}(\alpha+\eta)\chi_{\alpha+\eta}(z_1). \\
&= |H||H'| \left(\sum_{\eta \in \widehat{H'}^{\perp} \cap \widehat{H}} \widehat{A_{H'}^{g_1+z_1}}(0)\widehat{B_H^{g_2}}(\eta)\widehat{C_H^{g_3}}(\eta)\chi_{\eta}(z_1) \right. \\
&\qquad \left. + \sum_{\alpha \in \widehat{H'}, \eta \in \widehat{H'}^{\perp} \cap \widehat{H}, \alpha \neq 0} \widehat{A_{H'}^{g_1+z_1}}(\alpha)\widehat{B_H^{g_2}}(\alpha+\eta)\widehat{C_H^{g_3}}(\alpha+\eta)\chi_{\alpha+\eta}(z_1) \right) \\
&\geq |H||H'| \left(d_1 \sum_{\eta \in \widehat{H'}^{\perp} \cap \widehat{H}} \widehat{B_H^{g_2}}(\eta)\widehat{C_H^{g_3}}(\eta)\chi_{\eta}(z_1) \right. \\
&\qquad \left. - \frac{d_1d_2d_3}{4} \sum_{\alpha \in \widehat{H'}, \eta \in \widehat{H'}^{\perp} \cap \widehat{H}, \alpha \neq 0} \widehat{B_H^{g_2}}(\alpha+\eta)\widehat{C_H^{g_3}}(\alpha+\eta)\chi_{\alpha+\eta}(z_1) \right),
\end{aligned}
$$

where the last inequality uses the fact that the indicator function of A restricted to $H' + g_1 + z_1$ is $\frac{d_2d_3}{4}$-superregular on $H' + g_1 + z_1$, and that $\widehat{A_{H'}^{g_1+z_1}}(0) = d_1$. Using the Cauchy-Schwarz inequality, we get that the second term in the bracket is at least $-\frac{1}{4}d_1d_2d_3\sqrt{d_2d_3} \geq -\frac{1}{4}d_1d_2d_3$. Thus, if we have fewer than $\frac{1}{4}d_1d_2d_3|H||H'|$ triangles, we must have that the first term in the bracket is at most $\frac{1}{2}d_1d_2d_3$. This implies that,

$$
\sum_{\eta \in \widehat{H'}^{\perp} \cap \widehat{H}} \widehat{B_H^{g_2}}(\eta)\widehat{C_H^{g_3}}(\eta)\chi_{\eta}(z_1) \leq \frac{d_2d_3}{2}.
$$

We prove the following lemma, which allows us to deduce that H' shatters either B on $H + g_2$, or C on $H + g_3$. A proof has been included later in the section.

Lemma 4.6 (Fourier shattering) *Let $H' \preceq H \preceq G$. Given two functions f and g from H to $\mathbb{R}_{\geq 0}$ that satisfy*

$$
\sum_{\eta \in \widehat{H'}^{\perp} \cap \widehat{H}} \widehat{f}(\eta)\widehat{g}(\eta)\chi_{\eta}(z_1) \leq \frac{d_1d_2}{2}, \tag{2}
$$

for some positive d_1, d_2 and $z_1 \in H$; define the function $\overline{f}(v) \stackrel{\text{def}}{=} \mathbb{E}_{y\in H'}[f(v+y)]$ and $\overline{g}(v) \stackrel{\text{def}}{=} \mathbb{E}_{y\in H'}[g(v+y)]$, where $v \in H$. Then, either $\mathbb{P}_v\left[\overline{f}(v) < \frac{3}{4}d_1\right] > \frac{1}{20}$, or $\mathbb{P}_v\left[\overline{g}(v) < \frac{3}{4}d_2\right] > \frac{1}{20}$.

Assuming this lemma, and applying it to the functions $B_H^{g_2}$ and $C_H^{g_3}$, we deduce that for the functions $\overline{B_H^{g_2}}(v) \stackrel{\text{def}}{=} \mathbb{E}_{y\in H'}[B(v+y+g_2)] = \frac{|B\cap(H'+g_2+v)|}{|H'|}$ and $\overline{C_H^{g_3}}(v) \stackrel{\text{def}}{=} \mathbb{E}_{y\in H'}[C(v+y+g_3)] = \frac{|C\cap(H'+g_3+v)|}{|H'|}$, for $v \in H$,
we have either $\mathbb{P}_v\left[\overline{B_H^{g_2}}(v) < \frac{3}{4}d_2\right] > \frac{1}{20}$, or $\mathbb{P}_v\left[\overline{C_H^{g_3}}(v) < \frac{3}{4}d_3\right] > \frac{1}{20}$. This means that H' $(1/20, 3/4, \log|H/H'|)$-shatters either B on $H+g_2$ or C on $H+g_3$. ■

We are now ready to give a proof of the main lemma.

Proof: (*of Lemma 4.3*). Let $\rho \stackrel{\text{def}}{=} \frac{d_2 d_3}{4}$. Apply Corollary 3.3 with subgroup H, element g_1, regularity parameter ρ, density parameter $d_1/2$, and the set $A\cap(H+g_1)$, to partition $A\cap(H+g_1)$ into ρ-superregular parts. Let $\{(A_i, H_i, z_i)\}_{i=1}^t$ be the triples returned by Corollary 3.3 satisfying $|H/H_i| \le 2^{\log_{(1+\rho)}(2/d_1)}$, $A_i \subseteq H_i + z_i + g_1$, $|A_i| \ge d_1/2\,|H_i|$, and that the indicator function of A_i restricted to $H_i + z_i + g_1$ is ρ-superregular on $H_i + z_i + g_1$.
Fix a triple (A_i, H_i, z_i). If H_i $(1/20, 3/4, \log_2|H/H_i|)$-shatters A on $H+g_2$ or $H+g_3$, it also $(1/20, 3/4, \log_{1+\rho} 2/d_1)$-shatters it and we are done. Otherwise, applying Lemma 4.5 to the sets $A_i, A\cap(H+g_2)$, and $A\cap(H+g_3)$, on the cosets $H_i + z_i + g_1, H+g_2$, and $H+g_3$, respectively, there must be at least $\frac{1}{4}\frac{|A_i|}{|H_i|}d_2 d_3|H||H_i| = \frac{1}{4}d_2 d_3|A_i||H|$ triangles between $A_i, A\cap(H+g_2)$, and $A\cap(H+g_3)$.
Repeating the above argument for every triple, assume we do not find a subset H' that $(1/20, 3/4, \log_{1+\rho} 2/d_1)$-shatters A on at least one of $H+g_2$ or $H+g_3$. Then, since the sets $\{A_i\}_i$ are disjoint and for all i, $A_i \subseteq A$, the total number of triangles between $A\cap(H+g_1), A\cap(H+g_2)$, and $A\cap(H+g_3)$ is at least

$$\sum_{i=1}^{t}\frac{1}{4}d_2 d_3|A_i||H| = \frac{1}{4}d_2 d_3|H|\sum_{i=1}^{t}|A_i| \ge \frac{1}{8}d_1 d_2 d_3|H|^2,$$

where the last inequality follows because $\{A_i\}_{i=1}^t$ are disjoint subsets satisfying $|(A\cap(H+g_1))\setminus(A_1\cup\cdots\cup A_t)| \le \frac{d_1}{2}|H|$, and hence $\sum_i |A_i| \ge |A\cap(H+g_1)| - \frac{d_1}{2}|H| \ge \frac{d_1}{2}|H|$. This completes the proof. ■

We now give a proof of Lemma 4.6.

Proof: (*of Lemma 4.6*). We can simplify the left side of (2) as follows:

$$
\begin{aligned}
\sum_{\eta \in H'^{\perp} \cap \widehat{H}} \widehat{f}(\eta)\widehat{g}(\eta)\chi_\eta(z_1) &= \frac{1}{|H|^2} \sum_{x_1, x_2 \in H} \sum_{\eta \in H'^{\perp} \cap \widehat{H}} f(x_1)\chi_\eta(x_1) g(x_2)\chi_\eta(x_2)\chi_\eta(z_1) \\
&= \frac{1}{|H||H'|} \sum_{\substack{x_1, x_2 \in H \\ x_1 + x_2 + z_1 \in H'}} f(x_1) g(x_2) \\
&\text{(Since } x_1 + x_2 + z_1 \in H \text{ and } |H'||H'^{\perp} \cap \widehat{H}| = |H|\text{)} \\
&= \frac{1}{|H||H'|} \sum_{x_1 \in H, y_1 \in H'} f(x_1) g(x_1 + y_1 + z_1) \\
&= \frac{1}{|H||H'|} \sum_{y_1, y_2 \in H', v \in H \cap H'^{\perp}} f(v + y_2) g(v + y_2 + y_1 + z_1) \\
&= \frac{|H'|}{|H|} \sum_{v \in H \cap H'^{\perp}} \mathop{\mathbb{E}}_{y_2 \in H'} \left[f(v + y_2) \left(\mathop{\mathbb{E}}_{y_1 \in H'} [g(v + y_2 + y_1 + z_1)] \right) \right] \\
&= \frac{|H'|}{|H|} \sum_{v \in H \cap H'^{\perp}} \mathop{\mathbb{E}}_{y_2 \in H'} \left[f(v + y_2) \left(\mathop{\mathbb{E}}_{y_1 \in H'} [g(v + y_1 + z_1)] \right) \right] \\
&\text{(Since for } y_1 \text{ uniform over } H',\ y_1 + y_2 \text{ is also uniform over } H'\text{)} \\
&= \frac{|H'|}{|H|} \sum_{v \in H \cap H'^{\perp}} \mathop{\mathbb{E}}_{y_2 \in H'} [f(y_2 + v)] \mathop{\mathbb{E}}_{y_1 \in H'} [g(y_1 + v + z_1)] \\
&= \mathop{\mathbb{E}}_{v \in H \cap H'^{\perp}} \left[\overline{f}(v)\overline{g}(v + z_1) \right].
\end{aligned}
$$

Thus,

$$
\mathop{\mathbb{E}}_{v \in H \cap H'^{\perp}} \left[\overline{f}(v)\overline{g}(v + z_1) \right] \leq \frac{d_1 d_2}{2}.
$$

We now claim that either $\mathbb{P}_v \left[\overline{f}(v) < \frac{3}{4}d_1 \right] > \frac{1}{20}$, or $\mathbb{P}_v \left[\overline{g}(v) < \frac{3}{4}d_2 \right] > \frac{1}{20}$. This holds because otherwise, since $\overline{f}(v), \overline{g}(v) \geq 0$,

$$
\mathop{\mathbb{E}}_{v \in H \cap H'^{\perp}} \left[\overline{f}(v)\overline{g}(v + z_1) \right] \geq \left(1 - 2 \cdot \frac{1}{20} \right) \cdot \frac{3}{4}d_1 \cdot \frac{3}{4}d_2 = \frac{81}{160}d_1 d_2 > \frac{1}{2}d_1 d_2,
$$

which is a contradiction. ∎

5 Proof of the main theorem

The proof of Theorem 1.2 will follow by repeated applications of the following lemma.

Lemma 5.1 (Main Lemma) *Let $A \subseteq G$ be a set that is a union of εN disjoint triangles. Suppose we have partitioned G into cosets of a subgroup H, and let $T \stackrel{\text{def}}{=} |G/H|$. If A contains less than $\frac{\varepsilon^3}{64T^2}N^2$ triangles (not necessarily disjoint), then there is a subgroup $H' \preceq H$ such that:*

1. $\mathsf{Ent}_A(G, H') \geq \mathsf{Ent}_A(G, H) + \frac{\varepsilon}{3600}$.
2. $|G/H'| \leq c^T$, *where* $c = 2^{2\log_{1+\varepsilon^2/16}(4/\varepsilon)}$.

Proof: First, remove those elements from A which belong to cosets of H in which A has density less than $\varepsilon/2$. Let A' be what is left from A. Notice that the number of elements removed in this process is at most $\varepsilon N/2$, and hence A' contains at least $\varepsilon N/2$ disjoint triangles.

Let g_1, g_2, g_3, be a triangle in A', *i.e.*, $g_1 + g_2 + g_3 = 0$, and for $i = 1, 2, 3$, let d_i be the density of A' in the coset $H + g_i$. Note that $d_1, d_2, d_3 \geq \varepsilon/2$. Since A' contains at most $\frac{\varepsilon^3 N^2}{64T^2} \leq \frac{d_1 d_2 d_3}{8}|H|^2$ triangles, by Lemma 4.3, there is a subgroup $H_1 \preceq H$ such that H_1 $(1/20, 3/4, \log_{1+\varepsilon^2/16} 4/\varepsilon)$-shatters A' on at least one of $H + g_1$ or $H + g_2$.

For each coset $H + g$ that can be shattered, identify any one subgroup that $(1/20, 3/4, \log_{1+\varepsilon^2/16} 4/\varepsilon)$-shatters A' on $H + g$. Let H' be the intersection of all these subgroups. For a coset $H + g$ that is not shattered, we have $\mathsf{Ent}_A(H + g, H') - \mathsf{Ent}_A(H + g, H) \geq 0$ (Lemma 2.6, part 2). For every coset $H + g$ such that A' is shattered on $H + g$, observe that $A \cap (H + g) = A' \cap (H + g)$, and hence Lemma 4.2 implies,

$$\begin{aligned}
\mathsf{Ent}_A(H+g, H') - \mathsf{Ent}_A(H+g, H) &= \mathsf{Ent}_{A'}(H+g, H') - \mathsf{Ent}_{A'}(H+g, H) \\
&\geq \frac{1}{20}\left(1 - \frac{3}{4} + \mathsf{Ent}\left(\frac{3}{4}\right)\right) \cdot \frac{|A' \cap (H+g)|}{|H|} \\
&\geq \frac{1}{600}\frac{|A' \cap (H+g)|}{|H|}.
\end{aligned}$$

Since A' contains at least $\varepsilon N/2$ disjoint triangles, and at least one element from each of these triangles is contained in a coset that is shattered, at least $\varepsilon N/6$ of the elements of A' belong to a coset that has been shattered. Thus, averaging over all the cosets of H (using Remark 2.4),

$$\mathsf{Ent}_A(G, H') - \mathsf{Ent}_A(G, H) \geq \frac{1}{600} \cdot \frac{\varepsilon}{6} = \frac{\varepsilon}{3600}.$$

Note that $|H/H'| \leq 2^{T\log_{1+\varepsilon^2/16}(4/\varepsilon)}$, and hence,

$$|G/H'| \leq T2^{T\log_{1+\varepsilon^2/16}(4/\varepsilon)} \leq 2^{2T\log_{1+\varepsilon^2/16}(4/\varepsilon)} = c^T.$$

$\blacksquare$

Now we show how to deduce the main theorem using the above lemma.

Proof: (*of Theorem 1.2*). Let δ be such that δ^{-1} is a tower of twos of height $\Theta(\log {}^1/_\varepsilon)$ (the constant in Θ will be specified later). Assume for contradiction that A is ε-far from being triangle-free, and has less than δN^2 triangles. Thus, more than εN elements have to be removed from A to make it triangle-free. Consider a maximal set of disjoint triangles in A, of say $\varepsilon_0 N$ triangles, and let $A' \subseteq A$ be the union of these triangles. Thus, $|A'| = 3\varepsilon_0 N$. From now on, we will only work with A'. Since $A \backslash A'$ must be triangle-free, $3\varepsilon_0 N \geq \varepsilon N$, *i.e.*, $\varepsilon_0 \geq {}^\varepsilon/_3$.

Since A has at most δN triangles, A' also has at most δN triangles. Now, we repeatedly apply Lemma 5.1 in order to find successively finer partitions of G with increasing mean entropies. For $i = 0$, start with the trivial partition according to cosets of $H_0 \stackrel{\text{def}}{=} G$ for which,

1. the number of parts is $T_0 \stackrel{\text{def}}{=} |{}^G/_G| = 1$, and,
2. the mean entropy density of the partition is $\mathsf{Ent}_{A'}(G, H_0) = 3\varepsilon_0 \log 3\varepsilon_0$.

At step i, if $\delta N^2 \leq \frac{\varepsilon_0^3}{64T_i^2} N^2$, then we can apply Lemma 5.1 to refine the partition according to a subgroup $H_{i+1} \preceq H_i$, such that $\mathsf{Ent}_{A'}(G, H_{i+1}) \geq \mathsf{Ent}_{A'}(G, H_i) + \frac{\varepsilon_0}{3600}$. Moreover $T_{i+1} \leq 2^{2^{2T_i \log_{1+\varepsilon_0^2/16}({}^4/_{\varepsilon_0})}}$. Since δ^{-1} is a tower of twos of height $\Theta(\log {}^1/_\varepsilon) = \Omega(\log {}^1/_{\varepsilon_0})$, we can pick the constant inside the Θ large enough so that the condition $\delta \leq \frac{\varepsilon_0^3}{64T_i^2}$ is satisfied for all $i \leq t \stackrel{\text{def}}{=} \lceil 12000 \log {}^1/_{3\varepsilon_0} \rceil$. This implies $\mathsf{Ent}_{A'}(G, H_t) \geq 3\varepsilon_0 \log 3\varepsilon_0 + \frac{\varepsilon_0}{3600} \cdot 12000 \log \frac{1}{3\varepsilon_0} > 0$. However, we must always have $\mathsf{Ent}_{A'}(G, H_t) \leq 0$ (Lemma 2.6, part 1), and hence this is a contradiction. ■

Acknowledgements

The authors would like to thank Oded Regev and anonymous reviewers for their valuable comments.

References

[Alo02] Noga Alon, *Testing subgraphs in large graphs*, Random Structures Algorithms **21** (2002), no. 3-4, 359–370. MR 1945375 3

[BGRS12] Arnab Bhattacharyya, Elena Grigorescu, Prasad Raghavendra, and Asaf Shapira, *Testing odd-cycle-freeness in Boolean functions*, Combin. Probab. Comput. **21** (2012), no. 6, 835–855. MR 2981157 2

[BGS15] Arnab Bhattacharyya, Elena Grigorescu, and Asaf Shapira, *A unified framework for testing linear-invariant properties*, Random Structures Algorithms **46** (2015), no. 2, 232–260. MR 3302896 2

[BX10] Arnab Bhattacharyya and Ning Xie, *Lower bounds for testing triangle-freeness in boolean functions*, Proc. 21st Ann. ACM-SIAM Symp. on Discrete Algorithms, SIAM, 2010, pp. 87–98. 3

[CF13] David Conlon and Jacob Fox, *Graph removal lemmas*, Surveys in Combinatorics 2013 (2013). 1, 2, 3

[FK14] Hu Fu and Robert Kleinberg, *Improved lower bounds for testing triangle-freeness in Boolean functions via fast matrix multiplication*, Approximation, randomization, and combinatorial optimization, LIPIcs. Leibniz Int. Proc. Inform., vol. 28, Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2014, pp. 669–676. MR 3319023 3

[Fox11] Jacob Fox, *A new proof of the graph removal lemma.*, Annals of Mathematics **174** (2011), no. 1, 561–579. 1, 2, 3, 4, 5, 6, 10

[GN14a] Jim Geelen and Peter Nelson, *The critical number of dense triangle-free binary matroids*, arXiv preprint arXiv:1406.2588 (2014). 2

[GN14b] ______, *Odd circuits in dense binary matroids*, arXiv preprint arXiv:1403.1617 (2014). 2

[Gow97] Timothy Gowers, *Lower bounds of tower type for Szemerédi's uniformity lemma*, Geometric and Functional Analysis **7** (1997), no. 2, 322–337. 3

[Gow98] ______, *A new proof of Szemerédi's theorem for progressions of length four*, Geometric and Functional Analysis **8** (1998), no. 3, 529–551. 8

[Gre05] Ben Green, *A Szemerédi-type regularity lemma in abelian groups, with applications*, Geometric And Functional Analysis **15** (2005), 340–376, 10.1007/s00039-005-0509-8. 1, 2, 3, 5, 8

[KO09] Subhash Khot and Ryan O'Donnell, *SDP gaps and UGC-hardness for max-cut-gain*, Theory of Computing **5** (2009), no. 4, 83–117. 2

[KSV09] Daniel Král, Oriel Serra, and Lluís Vena, *A combinatorial proof of the removal lemma for groups*, Journal of Combinatorial Theory, Series A **116** (2009), no. 4, 971–978. 1, 3

[RCLG92] Fan R.K. Chung and Ronald L. Graham, *Quasi-random subsets of $\mathbb{Z}_n$*, Journal of Combinatorial Theory, Series A **61** (1992), 64–86. 8

[RS78] Imre Z. Ruzsa and Endre Szemerédi, *Triple systems with no six points carrying three triangles*, Combinatorics, Vol. II (Amsterdam), Colloq. Math. Soc. János Bolyai, vol. 18, North-Holland, 1978, pp. 939–945. MR 0519318 2

[Sze75] Endre Szemerédi, *On sets of integers containing no k elements in arithmetic progression*, Acta Arithmetica **27** (1975), 199–245. 2, 3